Coupling on the MHT - Supplementary material

February 20, 2012

S.1 Map of the GPS network

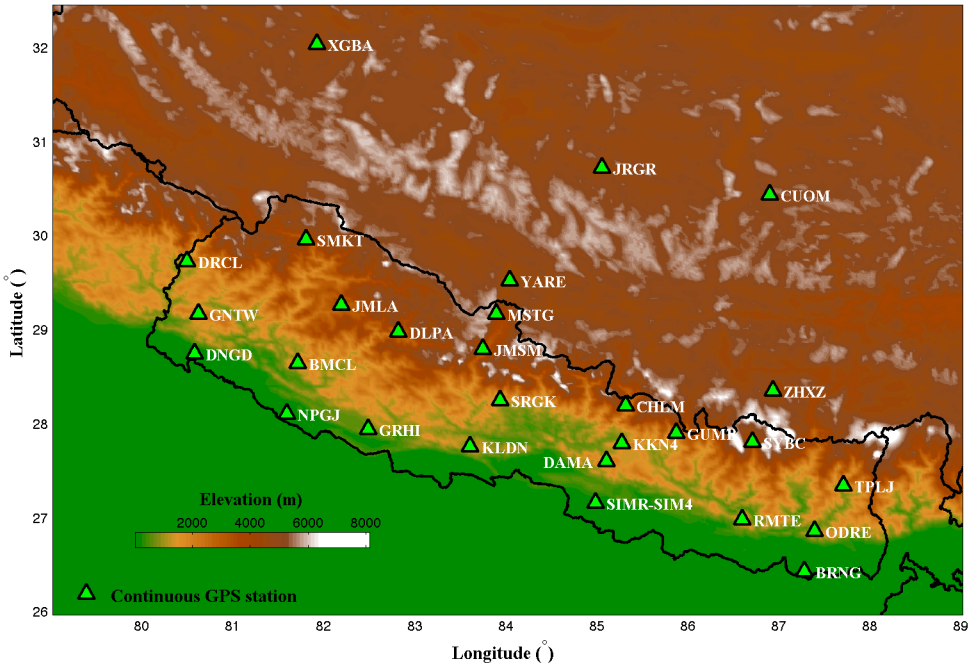


Figure S1: Map indicating the names and position of the continuous GPS stations used in this study to determine the pattern of coupling on the MHT.

S.2 Derivation of secular velocities from the GPS time series

S.2.1 Noise model for the inversion

Several studies of GPS time series have established that the daily estimates of GPS positions are temporally correlated (*Langbein and Johnson*, 1997; *Zhang et al.*, 1997; *Mao et al.*, 1999;

Williams, 2003a; *Williams et al.*, 2004). Assuming a purely white noise model is therefore incorrect and although it doesn't affect much the value of the final parameters inverted for, it results in a dramatic underestimation of their uncertainties. We thus add to the white noise in our GPS time series a component of colored noise, i.e. a noise that has a power spectrum of the form:

$$P_\kappa(f) \propto f^\kappa, \tag{S1}$$

where f is the temporal frequency, and κ is called the spectral index (*Mandelbrot and Van Ness*, 1968). The spectral index is estimated for each time series by first fitting equation (1) (see main paper) to the time series assuming a white noise, and computing a periodogram of the fit's residuals. The spectral index is then estimated by fitting to the power spectrum a combination of white and colored noise (figure S2):

$$P(f) = P_0 + P_c f^\kappa, \tag{S2}$$

where P_0 and P_c are the respective amplitudes of the white and the colored noise.

Once the spectral index is estimated, we build the covariance matrix of the data as the sum of white and colored noise covariance matrices. The relative amplitudes of both noises are estimated by a Maximum Likelihood Estimation (MLE) method (*Williams et al.*, 2004).

The covariance matrix for the white noise is the usual diagonal covariance matrix

$$\mathbf{C_w} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2),$$

where σ_i is the standard deviation of data point number i. The colored noise covariance matrix $\mathbf{C}_\kappa$ is built following an adaptation of the method described in *Williams* (2003a):

$$\mathbf{C}_\kappa = \Delta t_s \mathbf{T}\mathbf{T}^T, \tag{S3}$$

where Δt_s is the sampling interval (so $\Delta t_s = 1$ day for GPS time series), and the matrix $\mathbf{T}$ is defined as:

$$\mathbf{T} = \begin{pmatrix} \psi_0 & 0 & 0 & \ldots & 0 \\ \psi_1 & \psi_0 & 0 & \ldots & 0 \\ \psi_2 & \psi_1 & \psi_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_{n-1} & \psi_{n-2} & \psi_{n-3} & \ldots & \psi_0 \end{pmatrix}, \tag{S4}$$

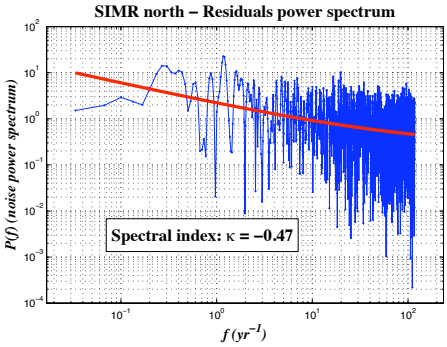
where the quantities ψ_n are defined by the recurrence formula:

$$\begin{cases} \psi_0 = 1 \\ \psi_{n+1} = \frac{n-\kappa/2}{n+1}\ \psi_n \end{cases} .$$

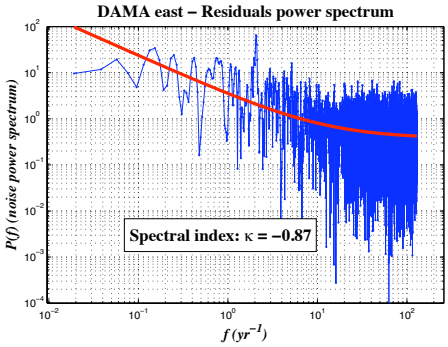
The rows and columns corresponding to times with no data are then removed from the covariance matrix.

The final data covariance matrix is given by

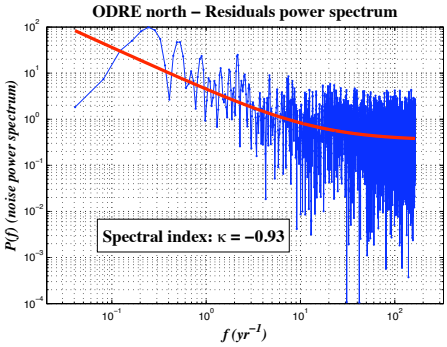
$$\mathbf{C_D} = a^2 \mathbf{C_w} + b^2 \mathbf{C}_\kappa, . \tag{S5}$$



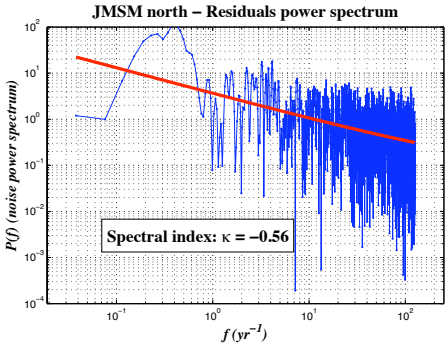
(a) SIMR, north component



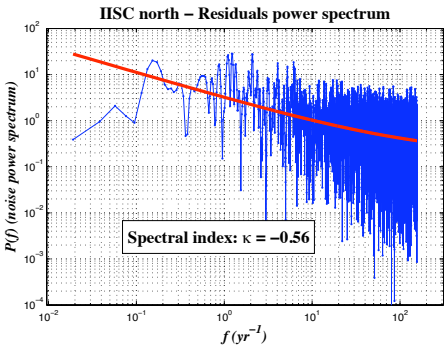
(b) DAMA, east component



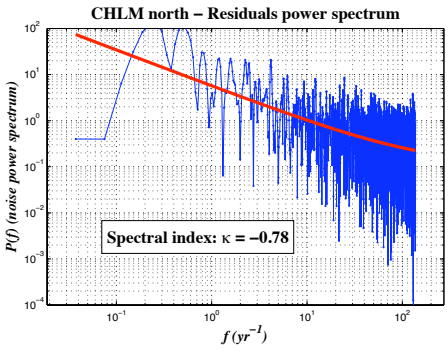
(c) ODRE, north component



(d) JMSM, north component



(e) IISC, north component



(f) CHLM, north component

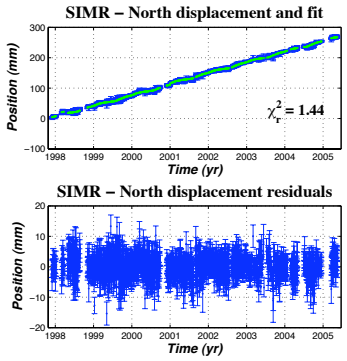
Figure S2: Power spectra of the noise (blue curve) for a sample of time series and fit assuming a combination of white noise and colored noise (red curve). The spectral index κ of the colored noise is indicated on each plot.

where a and b are the parameters to be estimated by MLE, measuring respectively the amplitude of white and colored noise. Assuming a Gaussian distribution of the uncertainties on GPS positions, the likelihood that has to be maximized with respect to a and b is then

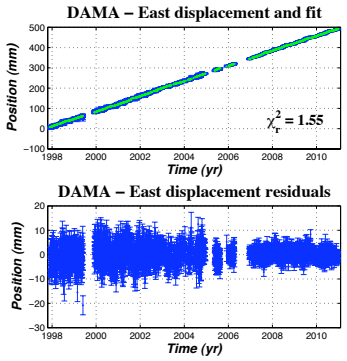
$$\text{likelihood}(\mathbf{C_D}) = \frac{1}{(2\pi)^{N/2}(\det \mathbf{C_D})^{1/2}} e^{-\frac{1}{2}\mathbf{r}^T {\mathbf{C_D}}^{-1}\mathbf{r}}, \tag{S6}$$

where $\mathbf{r}$ is vector of residuals of the fit and N is the number of daily GPS positions available.

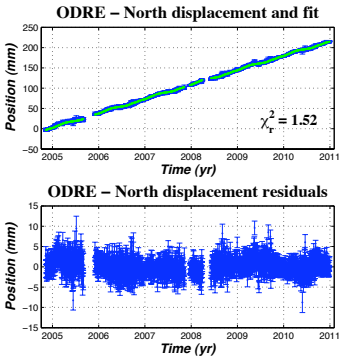
The fit and residuals on some time series are shown on figure S3.



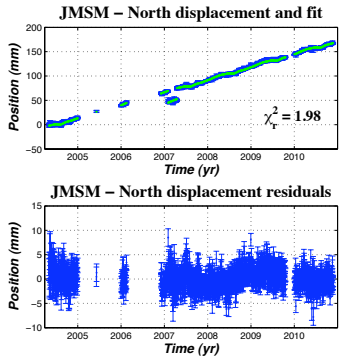
(a) SIMR, north component



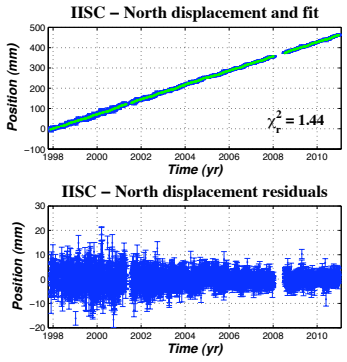
(b) DAMA, east component



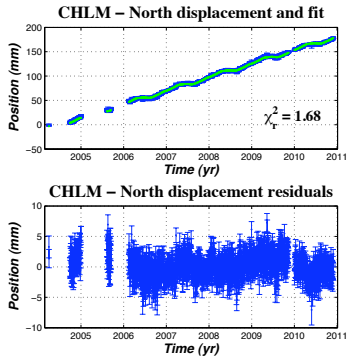
(c) ODRE, north component



(d) JMSM, north component



(e) IISC, north component



(f) CHLM, north component

Figure S3: Fits and residuals of the continuous GPS time series at some stations. For each time series, the upper plot shows the raw data (blue curve) and the fit (green curve) with equation (1). The value of the reduced chi square χ^2_r of each fit is indicated on the upper plot's lower right corner of the corresponding figure.

S.2.2 Uncertainties due to unmodeled steps in the time series

Steps in the time series can be of many different origins, being actually tectonic, environmental or coming from equipment malfunction, human error, etc. (*Williams*, 2003b). The

ones large enough to be detected are included in the model (equation (1)), but smaller ones remain unnoticed and affect the estimates of model parameters and their uncertainties. Therefore, those uncertainties have to be adjusted accordingly. For convenience, we will assume that those unmodeled steps account for all the errors on the model.

Those steps are assumed to happen at a frequency ν, and to have a random Gaussian amplitude $\mathcal{N}(0, \sigma_x^2)$. The standard deviation on the secular velocity due to those steps is then (*Williams*, 2003b)

$$\sigma_v = \frac{\sigma_x \sqrt{\nu}}{\sqrt{T}}, \tag{S7}$$

where T is the length of the time series.

In the case of our GPS time series, the amplitude of the steps that were actually detected was always greater than 1.5 time the median value of the uncertainties on the daily positions in the time series. We hence take $\sigma_x = \langle \sigma_D \rangle$, where $\langle . \rangle$ denotes the median value and σ_D is the uncertainty on daily positions of the time series.

We estimate ν through the following considerations. First, the steps that were large enough to be detected in the time series happened on average once every 5 years. Assuming that the smaller the steps are, the more frequent they would be, the value for ν should be greater than 1/5. On the other hand, a value of ν overestimated ($\nu > 1$ in this case) results in larger uncertainties on the secular velocity, and eventually leads to values of a reduced chi square smaller than 1 when one fits the Euler pole of the Indian plate in the ITRF 2005 reference frame (see section 3.3), indicating that the uncertainties on the GPS velocities are probably overestimated. As a result, we chose a value of $\nu = 1/3$, which gives the final formula for the uncertainties on the secular velocity due to unmodeled steps:

$$\sigma_v = \frac{\langle \sigma_D \rangle}{\sqrt{3T}}. \tag{S8}$$

The velocities and corresponding uncertainties that we obtain at the GPS stations used in this paper as well as at the DORIS stations COLA and EVEB are given in table S2.

S.3 Slip resolution and Laplacian

The result of our inversion should be assessed in view of its resolution. This information is contained in the resolution matrix:

$$\mathbf{R} = \left(\mathbf{G}^T {\mathbf{C_d}}^{-1} \mathbf{G}^T + \mathbf{\Lambda}^T \mathbf{\Lambda} \right)^{-1} \mathbf{G}^T {\mathbf{C_d}}^{-1} \mathbf{G}^T, \tag{S9}$$

where $\mathbf{G}$ is the Green's matrix defined in equation (5) from the main paper, $\mathbf{C_d}$ is the data covariance matrix and $\mathbf{\Lambda}$ is the Laplacian matrix. The diagonal of $\mathbf{R}$ tells how well the slip value on each patch can be retrieved by the inversion. However, it doesn't express how each patch correlates with its neighbors. This information is contained in each of the individual columns of $\mathbf{R}$: column number i is the vector of parameters (i.e. the slip on each patch) returned by the inversion from an input dataset corresponding to a unit slip on patch i and no slip on other patches. Usually, what the inversion returns is slip on a more or less

spread area centered on patch i. The characteristic size of this area is estimated by fitting a bell curve to the slip on the patches as a function of distance to patch i (*Lohman*, 2004), and taking the standard deviation of that bell curve. Namely, for each patch i, we find the distance w_i that minimizes the quantity:

$$\chi_i^2 = \sum_{j=1}^{N_\mathrm{p}} \left(\left| \frac{R_{ji}}{R_{ii}} \right| - e^{-\frac{d_{ij}^2}{2w_i^2}} \right)^2 , \tag{S10}$$

where N_p is the number of patches on the fault, R_{ji} is the value of the coefficient (j, i) of the resolution matrix $\mathbf{R}$ (row j and column i), and d_{ij} is the distance between patches i and j.

This idea of an estimate of the resolution scale on each patch is also used in order to more efficiently smooth our model by weighting the Laplacian according to the resolution on each patch. Since the Laplacian matrix is not yet available (this is what we try to determine), we compute a first resolution matrix using the Moore-Penrose pseudoinverse matrix (*Aster et al.*, 2005), keeping only the singular values larger than 10% of the maximum one. We then compute how far each patch correlates with its neighbors with the method previously described applied to this resolution matrix. Finally, each line of the Laplacian matrix is weighted by the decimal logarithm of the resolution size on the corresponding patch.

S.4 Supplementary figures on the pattern of coupling on the MHT

S.4.1 Laplacian smoothing

On figure S4 we test how different values of the Laplacian smoothing affect the estimate of the moment deficit accumulated every year. Weights assigned to the Laplacian too small

Station ID	Site name	Latitude (°N)	Longitude (°E)	Elevation (m)
BAN2	Bangalore	13.03431	77.51161	832
DGAR	Diego Garcia	-7.26968	72.37024	-65
GUAO	Guao	43.47111	87.17731	2029
GUAM	Guam Observatory	13.58933	144.86836	202
HYDE	Hyderabad	17.41726	78.55087	442
IISC	Indian Inst. Science	13.02117	77.57038	844
KUNM	Kunming	25.02954	102.79712	1986
LHAS	Lhasa	29.65734	91.10399	3625
LHAZ	Lhasa2	29.67533	91.10403	3625
POL2	Poligan IVTAN 2	42.67977	74.69427	1714
SELE	Selezaschita	43.17873	77.01690	1342
TAIW	Taipei	25.02133	121.53654	44
URUM	Urumqi	43.80795	87.60067	859
WUHN	Wuhan	30.53165	114.35726	26

Table S1: List of IGS sites included in the daily regional processing.

Station	lon ($^\circ E$)	lat ($^\circ N$)	Velocities in ITRF05 (mm/yr)			Time of operation	
			V_e	V_n	V_u	Init.	End
DAMA	85.1077	27.6081	36.73 ± 0.45	34.21 ± 0.33	1.46 ± 1.37	Nov 1997	current
GUMB	85.8775	27.9098	35.88 ± 0.44	28.2 ± 0.34	5.7 ± 1.19	Nov 1997	current
SIMR	84.9844	27.1646	37.05 ± 1.03	34.82 ± 0.58	2.68 ± 2.72	Nov 1997	Apr. 2005
BRNG	87.2813	26.4387	37.7 ± 2.83	36.15 ± 3.49	-2.2 ± 3.55	Mar 2004	May 2009
BRN2	87.272	26.5197	38.46 ± 1.34	33.61 ± 1.82	8.13 ± 5.48	May 2009	current
CHLM	85.3154	28.2054	36.65 ± 0.48	27.55 ± 0.35	4.36 ± 1.2	Mar 2004	current
JMSM	83.7467	28.8044	34.33 ± 0.54	26.28 ± 0.39	3.28 ± 1.36	Ma. 2004	current
KKN4	85.2788	27.8008	36.12 ± 0.45	32.27 ± 0.4	1.13 ± 1.23	Jan 2004	current
KLDN	83.6119	27.7669	35.66 ± 0.43	34.64 ± 0.36	1.69 ± 1.17	Apr 2004	current
MSTG	83.8946	29.1789	34.68 ± 1.3	24.19 ± 1.05	5.79 ± 3.7	Apr 2004	Sept 2004[a]
MST2	83.953	29.1778	31.64 ± 1.59	23.73 ± 1.13	1.81 ± 4.09	Oct 2009	current
ODRE	87.3921	26.8662	38.71 ± 0.52	35.44 ± 0.37	-3.02 ± 1.5	Mar 2004	current
SIM4	84.99	27.17	37.13 ± 0.63	35.32 ± 0.5	-1.4 ± 1.8	Mar 2004	current
SRGK	83.9358	28.2603	35.62 ± 1.03	30.65 ± 0.9	4.52 ± 3.01	Mar 2005	Feb 2007
TPLJ	87.71	27.35	37.98 ± 0.55	31.24 ± 0.32	1.25 ± 1.03	Mar 2004	current
BMCL	81.7144	28.6558	34.51 ± 0.93	33.5 ± 0.47	0.47 ± 2.29	Mar 2007	current
DLPA	82.8204	28.9853	34.85 ± 0.63	25 ± 0.51	1.23 ± 1.68	May 2007	current
GRHI	82.4914	27.9509	35.09 ± 0.62	32.27 ± 0.54	3.97 ± 1.64	May 2007	current
JMLA	82.1923	29.2742	32.04 ± 0.76	26.14 ± 0.45	2.1 ± 1.6	May 2007	current
NPGJ	81.5953	28.1172	35.22 ± 0.67	39.04 ± 0.74	-0.16 ± 1.63	May 2007	current
BYNA	81.2007	29.4742	31.5 ± 1.38	26 ± 0.74	1.69 ± 2.39	May 2008	current
DNGD	80.5818	28.7545	35.29 ± 0.67	30.14 ± 0.85	-0.48 ± 1.8	May 2008	current
DRCL	80.5009	29.7338	31.41 ± 0.72	29.78 ± 1.2	2.64 ± 2.34	Mar 2008	current
GNTW	80.6262	29.1765	33.57 ± 0.67	33.68 ± 0.63	0.72 ± 3.15	Apr 2008	current
RMJT	86.55	27.3051	35.08 ± 1.62	32.58 ± 1.48	-1.04 ± 4.36	Oct 2008	current
RMTE	86.5971	26.991	35.86 ± 0.78	33.49 ± 0.69	1.44 ± 2.02	Sep 2008	current
SMKT	81.8065	29.9694	29.82 ± 0.71	21.84 ± 0.52	3.73 ± 1.83	May 2008	current
SYBC	86.7125	27.8142	35.55 ± 1.18	25.93 ± 0.96	7.14 ± 2.88	Oct 2008	current
CUOM	86.9039	30.4451	40.52 ± 0.59	20.38 ± 0.52	0.3 ± 2.05	Oct 2006	current
JRGR	85.0568	30.7286	36.11 ± 0.59	20.39 ± 0.5	3.03 ± 1.57	Mar 2007	current
XGBA	81.9259	32.0469	29.19 ± 1.55	18.25 ± 1.43	1.47 ± 4.25	Mar 2007	Sep 2007[b]
YARE	84.0431	29.5344	35.56 ± 0.68	24.14 ± 0.53	3.56 ± 1.86	Oct 2006	current
ZHXZ	86.9396	28.3569	37.86 ± 0.54	24.36 ± 0.45	1.87 ± 1.88	Oct 2006	current
MALD	73.526	4.189	43.35 ± 0.65	34.9 ± 0.49	-5.1 ± 1.93	Jul 1999	May 2006
HYDE	78.551	17.417	39.24 ± 0.49	35.24 ± 0.36	0.48 ± 1.18	Sept 2002	current
IISC	77.5704	13.0212	41.74 ± 0.47	35.06 ± 0.35	-0.05 ± 1.26	Oct 1997	current
COLA	79.8741	6.892	44.54 ± 4.8	35.33 ± 2.99	0.44 ± 3.78	Jan 1993	Sep 2004
EVEB	86.8131	27.9581	37.08 ± 4.4	25.32 ± 3.17	2.06 ± 3.64	May 1993	current

Table S2: Estimates of the secular velocity at the continuous GPS stations in ITRF 2005 and dates of operation of each station. The uncertainties on the velocities indicated are the 1-σ uncertainties. See text for details on the derivation of those quantities. Gaps in the time series are not unfrequent, and one should keep in mind that they are not indicated in this table.

[a] A 2-day campaign measurement has also been done with a different antenna on the station's monument in October 2009.

[b] 4 additional points in May 2009 made the positions at this station exploitable.

($\lambda < 0.8$) lead to models featuring locked patches only underneath data points, right next to creeping patches. Besides being unphysical and resulting in very high reduced chi squares, such models are highly dependent on the data spatial distribution and must then be rejected. A smoothing too large ($\lambda > 5$) tends to lead to a fault locked further at depth, and with a very smooth locked-creeping transition, which doesn't fit the data anymore (reduced $\chi^2 > 3$ on figure S4). Within the range of Laplacian weight $0.8 < \lambda < 5$, the moment deficit accumulated each year remains within the uncertainties determined by the inversion.

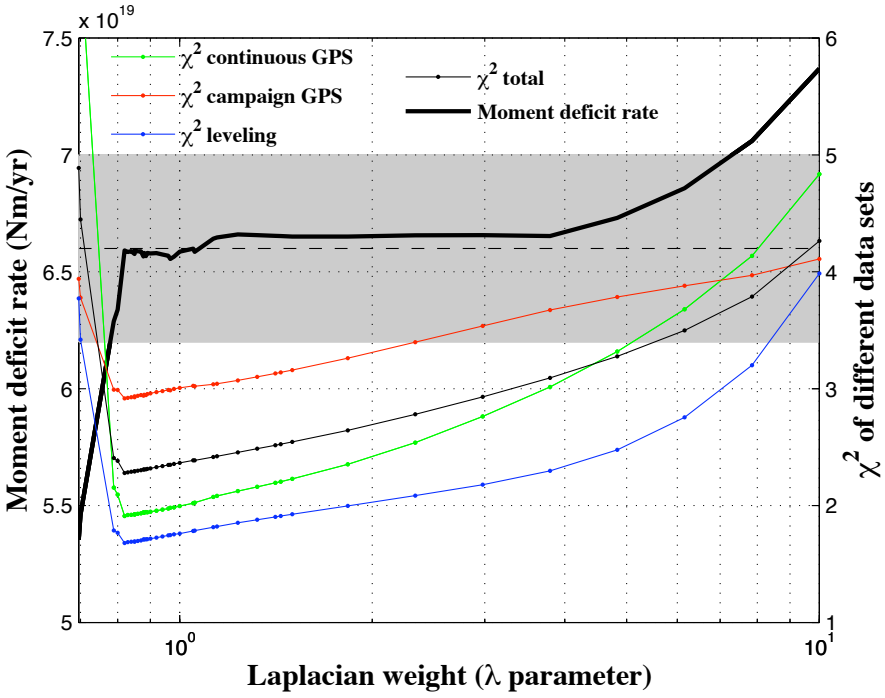


Figure S4: Variation of the χ^2 of the fit and of the moment deficit rate for different values of the weight attributed to the Laplacian in the inversion. The black curve shows the moment deficit accumulated every year as a function of the weight attributed to the Laplacian. The dashed black line and grey shaded area represent the rate of moment deficit with uncertainties derived in this study, i.e $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr. The green, red and blue curves respectively represent the value of the χ^2 of the fit to the continuous GPS, campaign and leveling data.

S.4.2 Direction of extension of the Tibetan plateau

Figure S5 shows the sensitivity of the long term velocity and the moment accumulation rate estimated in this study to the direction chosen for the extension of the Tibetan Plateau. The direction N98.2E has been chosen because it is the one that affects the least the estimates of the long term velocities (it is the 'most perpendicular' direction to those velocities, i.e. it is the direction onto which the sum of the projections of the East and West long term velocities

reaches a minimum). But there is no real reason to prevent this direction from varying by a few degrees from the N98.2E azimuth. Figure S5 shows that even by changing this direction by 10°, the final values of the parameters remain within their estimated uncertainties.

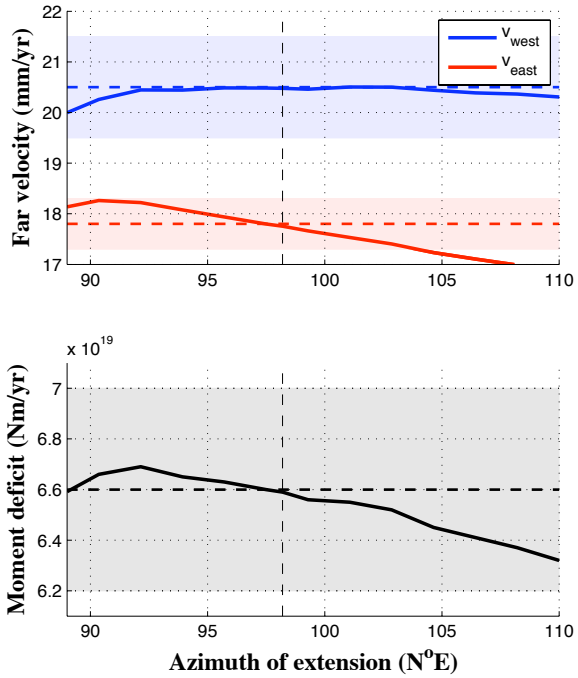
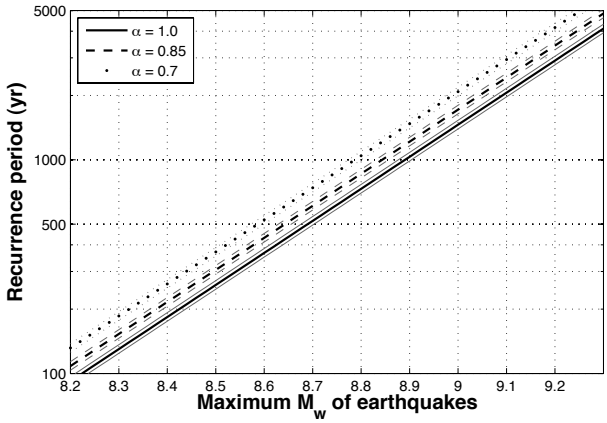


Figure S5: Impact of the azimuth selected for the extension of the Tibetan plateau on the long term East and West velocities (upper plot) and the moment deficit rate (lower plot). The solid thick lines represent the values of the parameters with respect the azimuth, the horizontal dashed lines and filled area of corresponding colors are the values with 1-σ uncertainties that we retained in this study (corresponding to an azimuth of N98.2E): $V_e = 17.8 \pm 0.5$ mm/yr, $V_w = 20.5 \pm 1$ mm/yr and $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr.

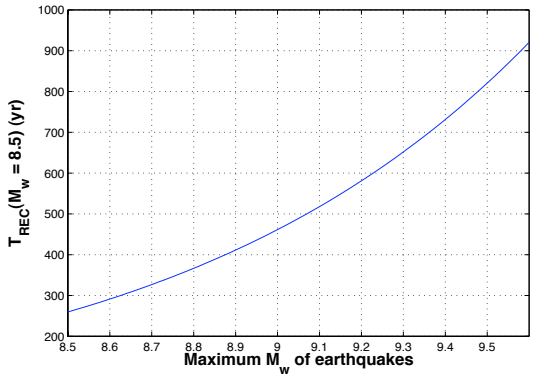
S.4.3 Recurrence time of large eartquakes

Assuming that the moment deficit of $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr computed in the main paper was released through earthquakes following a Gutenberg-Richter distribution up to a maximum magnitude, the recurrence time of those largest earthquakes (corresponding to those largest magnitudes) is plotted on figure S6(a). The black lines (solid, dashed and dotted) correspond to $\dot{\mathcal{M}}_0 = 6.6 \times 10^{19}$ Nm/yr, with different percentages of this moment deficit being released seismically, while the grey surrounding lines show the extent corresponding to the uncertainties on $\dot{\mathcal{M}}_0$. This plot shows that earthquakes as large as the 1950 Assam earthquake, whose moment magnitude is estimated at $M_w \sim 8.5$ (*Ambraseys and Douglas*, 2004; *Chen and Molnar*, 1977), could happen as often as once every 270 years within the borders of Nepal. As far as frequency is concerned, this would be the worst case

scenario where all the moment deficit accumulated was released seismically in earthquakes whose magnitude wouldn't exceed 8.5. However, too many parameters remain unknown to make any accurate estimation on the return period of major earthquakes. Should the actual b-value of the seismicity distribution in Nepal slightly differ from 1, equation (8) shows that those estimates would be significantly affected. Another unknown parameter is the largest possible earthquake magnitude in Nepal which has a paramount effect, as shown on figure S6(b). Indeed if the seismicity on the MHT doesn't go beyond those $M_w \sim 8.5$ earthquakes, they would indeed have a period of return of about 270 years. But if we assume that the MHT can produce earthquakes up to $M_w \sim 9.2$, then the return period of $M_w \geq 8.5$ earthquakes would become of the order of 600 years.



(a) Recurrence time of the largest possible earthquakes in Nepal assuming a release of a proportion α of the accumulated moment by a seismicity following a Gutenberg-Richter distribution with $b = 1$.



(b) Reccurence time of $M_w \geq 8.5$ earthquakes as a function of the largest possible earthquakes happening in Nepal, for a moment accumulation of $\dot{\mathcal{M}}_0 = 6.6 \times 10^{19}$ Nm/yr released entirely seismically.

Figure S6: Estimations on the recurrence time of earthquakes.

References

Ambraseys, J. J., and J. Douglas, Magnitude calibration of north indian earthquakes, *Geophysical Journal International*, *159*, 165–206, 2004.

Aster, R. C., C. H. Thurber, and B. Borchers, *Parameter Estimation and Inverse Problems*, Academic Press, 2005.

Chen, W.-P., and P. Molnar, Seismic moments of major earthquakes and the average rate of slip in central asia, *Journal of Geophysical Research*, *82(20)*, 29452969, 1977.

Langbein, J., and H. Johnson, Correlated errors in geodetic time series: Implications for time-dependent deformation, *Journal of Geophysical Research*, *102*, 591–603, 1997.

Lohman, R. B., The inversion of geodetic data for earthquake parameters, Ph.D. thesis, California Institute of Technology, 2004.

Mandelbrot, B., and J. Van Ness, Fractional brownian motions, fractional noises, and applications, *SIAM Rev.*, *10*, 422–439, 1968.

Mao, A., C. G. A. Harrison, and T. H. Dixon, Noise in gps coordinate time series, *Journal of Geophysical Research*, *104*, 2797–2816, 1999.

Williams, S. D., Y. Bock, P. Fang, P. Jamason, R. M. Nikolaidis, L. Prawirodirdjo, M. Miller, and D. J. Johnson, Error analysis of continuous gps position time series, *Journal of Geophysical Research*, *109*, 2004.

Williams, S. D. P., The effect of coloured noise on the uncertainties of rates estimated from geodetic time series, *Journal of Geodesy*, *76*, 483–494, 2003a.

Williams, S. D. P., Offsets in global positioning system time series, *Journal of Geophysical Research*, *108*, doi:10.1029/2002JB002,156, 2003b.

Zhang, J., Y. Bock, H. Johnson, P. Fang, S. Williams, J. Genrich, S. Wdowinski, and J. Behr, Southern california permanent gps geodetic array: Error analysis of daily position estimates and site velocities, *J. Geophys. Res.*, *102*, 18,035–18,055, 1997.